Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Reliability of atomic displacement parameters in protein crystal structures

Oliviero Carugo[a,b]* and Patrick Argos[a]

[a]European Molecular Biology Laboratory, Meyerhofstrasse 1, 69117 Heidelberg, Germany, and [b]Department of General Chemistry of the University, Via Taramelli 12, 27100 Pavia, Italy

Correspondence e-mail: carugo@embl-heidelberg.de

Received 10 July 1998
Accepted 11 September 1998

Mean standard errors in atomic displacement parameters (ADPs) resulting from protein crystal structure determinations are estimated by comparing the ADPs of protein-chain pairs of identical sequence within the same crystal or within different crystals displaying the same or different space groups. The estimated ADP standard errors increase nearly linearly as the resolution decreases – an unexpected result given the nonlinear dependence of the resolution on the amount of diffraction data. The estimated ADP standard errors are larger for side-chain and solvent-exposed atoms than for main-chain and buried atoms and, surprisingly, are also larger for residues in the helical secondary structure relative to other local backbone conformations. The results allow an estimate of the influence of crystallographic refinement restraints on ADP standard errors. Such corrections should be applied when comparing different protein structures.

1. Introduction

Atomic displacement parameters (ADPs, often referred to as thermal or *B* factors) are given increasing significance in protein crystallography since they provide information on the flexibility of main-chain and side-chain atoms when anisotropically (Dauter *et al.*, 1995; Longhi *et al.*, 1997; Harata *et al.*, 1998) or isotropically (Carugo & Argos, 1997, 1998; Luedemann *et al.*, 1997) refined. The ADPs are also useful to infer the precision of atomic positional coordinates (Murshudov & Dodson, 1997; Cruickshank, 1996) and to provide a statistical basis for comparing different structural results (Carugo & Eisenhaber, 1997; Peters-Libeu & Adman, 1997). The best predictions of the flexibility of a polypeptide chain have been based on analyses of ADPs in known structure test sets (Karplus & Schulz, 1985; Bhaskaran & Ponnuswamy, 1988; Vihinen *et al.*, 1994; Parthasarathy & Murthy, 1997, 1998). Unfortunately, ADP mean values and variances often differ greatly amongst various protein structure determinations, since absolute ADP values not only depend on various physical phenomena such as authentic oscillations about stable conformations or static, dynamic and lattice disorders, but also depend on differences in structure-refinement methods and stages (Ringe & Petsko, 1986).

Various attempts have been made to overcome ADP variances. The minimum-function method equalizes the minimum ADP values found in two protein structures (Frauenfelder & Petsko, 1980; Ringe & Petsko, 1986). The ADPs of each protein structure can be constrained to a common mean and a common variance: for example,

Table 1
Number of type 1, 2, and 3 structure pairs at various resolution ranges compared in the present paper.

Resolution range (Å)	Number of structure pairs			
	Type 1	Type 2	Type 3	Total
1.4–1.5	63	92	51	206
1.6–1.7	153	205	104	462
1.8–1.9	711	376	348	1435
2.0–2.1	933	830	716	2479
2.2–2.3	633	232	250	1115
2.4–2.5	787	599	121	1507
2.6–2.7	462	83	160	705
2.8–2.9	369	125	204	698
3.0–3.1	81	28	79	189
3.2–3.3	87	57	38	182
3.4–3.5	79	0	11	90

$ADP_{norm} = (ADP + D)/(\langle ADP\rangle + D)$, where $\langle ADP\rangle$ is the mean value for a given protein and D is an empiric parameter, to yield normalized ADP values (ADP_{norm}) with mean 1.0 and root-mean-square deviation 0.3 (Karplus & Schulz, 1985; Vihinen *et al.*, 1994). Another commonly used expression is $ADP_{norm} = (ADP - \langle ADP\rangle)/ADP_s$ (where ADP_s is the standard deviation of the ADP distribution) yielding, for each protein structure, zero mean and unit variance (Parthasarathy & Murthy, 1997, 1998; Carugo & Argos, 1997, 1998).

The need for ADP normalization results from a lack of fundamental understanding of their etiology. The main drawback of the minimum-function method is that the ADP variances within each single structure are not considered; they may be very different according to the severity of the ADP restraint applied during refinement. Any normalization imposing the same mean and variance to ADPs from different structures yields values insensitive to particular spatial conditions in a given protein structure. It is preferable to use experimentally derived ADPs, along with a reliable estimate of their accuracy. In principle, this is possible with full-matrix least-squares refinement based on simultaneous use of ADPs and atom coordinates; nevertheless, the difficulty in applying proper weights for refinement restraints on ADPs and atomic coordinates remains. Furthermore, such restraints may analytically differ amongst various software packages and may be perceived differently by crystallographers. In the present communication, an estimation of ADP reliability based on normal probability-plot analyses over known protein crystal structures is provided.

2. Methods

The Protein Data Bank (PDB; Bernstein *et al.*, 1977) was scanned to search for pairs of identical polypeptide sequences at least 50 residues in length and belonging to the same PDB file (referred to as type 1 pairs and generally consisting of multimeric proteins), or to different PDB files with the same (type 2) or different space group (type 3). Sequences were taken from the ATOM records in the PDB files. Structure pairs were considered only if their crystallographic resolution differed by no more than 0.1 Å. Crystal structures refined by multi-conformer or related techniques (Burling & Brünger, 1994) were rejected, as were those refined by inclusion of non-crystallographic symmetries, by visual inspection of the PDB files. Table 1 summarizes the number of structure pairs compared according to crystallographic resolution, with distribution centred around 2.1 Å, which is in good agreement with the mean PDB value of 2.19 Å and variance 0.50 Å. The relatively large number of structure pairs shown in Table 1 is not surprising given the combinatorial nature of the pair definition; for example, one structure with three monomers in the asymmetric unit provides three pairs and four trimeric structures define 66 pairs.

Secondary-structural assignments of each protein residue were performed with *STRIDE* (Frishman & Argos, 1995), where residues were classified as helical (α, 3_{10} or π), strand (extended conformation), turn (generally short turns explicitly designated by *STRIDE*) or coil (all remaining residues). Solvent-accessible areas were computed for each protein residue with *ASC* (Eisenhaber *et al.*, 1995) using a water-probe radius of 1.25 Å, as recommended by Hubbard & Argos (1995). Residues were considered exposed if the solvent-accessible area was at least 50% of that of a reference value obtained by averaging the maximal solvent-accessible area for the residue type within 137 unique protein structures selected with *OBSTRUCT* [structure pairs with maximal sequence identity 20%, resolution better than or equal to 1.8 Å and minimum residue number 50 (Heringa *et al.*, 1992)].

All individual ADP values were taken in B units (and are expressed as such in this work) and were corrected if reported otherwise (typically U) in the PDB files. Assuming that corresponding ADP values are statistically identical in a pair of comparable structures, normal probability plots can be used to estimate their mean standard error (Abrahams & Keve, 1971; Hamilton, 1972). Normal probability plots were originally designed to evaluate the similarity between pairs of experimental data sets. Given two measurement sets X and Y, the weighted differences between the equivalent data x_i and y_i are computed as $(x_i - y_i)/(\sigma_{xi}^2 + \sigma_{yi}^2)^{1/2}$, where σ_{xi} and σ_{yi} are the standard errors associated with x_i and y_i, respectively. These weighted differences are normally distributed if the two data sets are statistically identical. This is verified by plotting, after sorting in order of increasing magnitude, the observed weighted differences *versus* the expected weighted differences (D_e) computed as $|(n - 2i + 1)/n|$, where $i = 1$ to n for a data set of n elements and with internal sign positive for $i > n/2$ and negative for $i < n/2$. A regression line with unit slope and zero intercept results. Different slopes indicate whether the two data sets are statistically different or whether the standard errors associated with each single measurement are over-estimated or underestimated. Deviations from linearity suggest that systematic errors affect the quality of the data. In the present work, normal probability plots were used to estimate ADP mean standard error values, assuming the two data sets (*i.e.*, the ADPs of equivalent atoms in two structures) to be statistically identical. The unweighted observed differences $D_o = x_i - y_i$ were sorted in order of increasing magnitude and plotted against the expected weighted differences D_e. The

slope a of the regression line $D_o = aD_e$ was then computed and assumed to be related to the mean standard error of the data (σ) by the equation $a = 2^{1/2}\sigma$. For example, the nine atoms of residues Gly114 and Ala115 of the crystal structure of the cathepsin B (PDB file 1CPJ; Jia *et al.*, 1995) have refined ADP values of 18.16, 43.35, 30.66, 33.26, 48.76, 18.48, 49.12, 38.19 and 60.77 Å^2 in chain A and 27.07, 28.04, 56.90, 38.62, 40.83, 2.42, 60.54, 28.21 and 64.25 Å^2 in chain B. The D_o values sorted in order of increasing magnitude are −26.24, −11.42, −8.91, −5.36, −3.48, 7.93, 9.98, 15.38 and 16.06 Å^2 and are compared with the sorted D_e values −0.889, −0.667, −0.444, −0.222, 0.000, 0.222, 0.444, 0.667 and 0.889 (Fig. 1). By least-squares regression it is found that $D_o = 22.5D_e$ and thus the mean estimated standard error σ associated with the ADPs is 16 Å^2. A plot of $D_o/(2\sigma^2)^{1/2}$ *versus* D_e results in an approximate line of unit slope and zero intercept (Fig. 1). In the normal probability plot analysis performed here, equivalent residues with solvent-accessible areas differing by more than 20% were disregarded since their local conformations are highly different (5% of the total). The structure pairs for which the linear correlation coefficient between D_o and D_e was lower than 0.85 were also disregarded (19% of the total) since large deviations from linearity could be caused by systematic errors resulting, for example, from a structure pair where only one member contains one or more large ligands burying several residues.

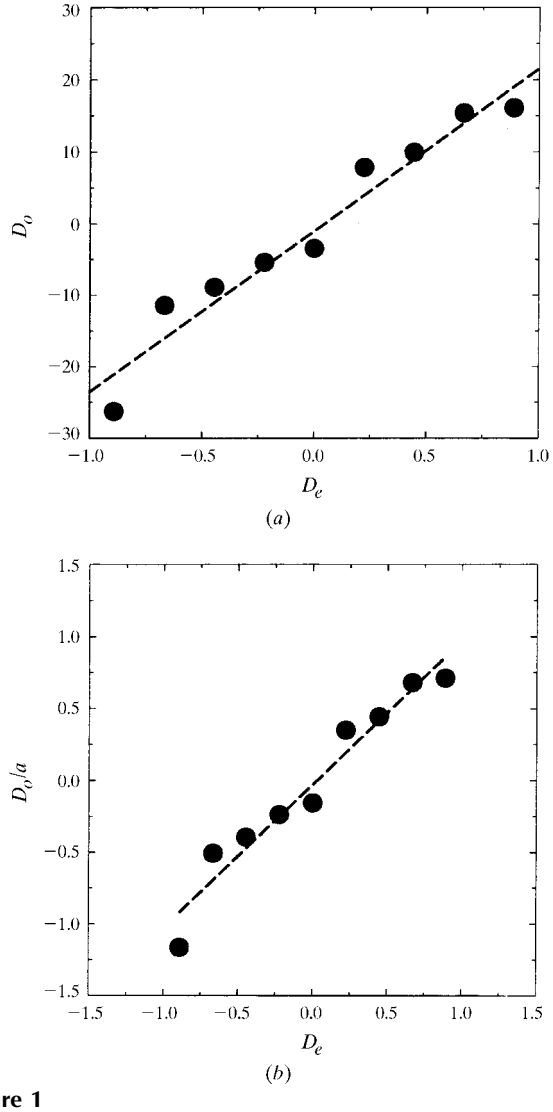

Figure 1
(*a*) Typical example of a normal probability plot where the slope a (22.5 Å^2) suggests an estimated ADP standard error of 16 Å^2. Data was taken from comparison of the Gly114 and Ala115 atoms of the A and B chains of the PDB file 1CPJ. A circle is plotted for each atom pair. (*b*) The same normal probability plot where the differences D_o are weighted with the reciprocal of the slope a. A regression line of unit slope and zero intercept results.

3. Results and discussion

Figs. 2 and 3 show the respective dependence of the estimated ADP standard errors on crystallographic resolution for main-chain and side-chain atoms (Fig. 2) and for solvent-exposed and buried atoms (Fig. 3). Clearly, the estimated ADP standard errors tend to increase as resolution decreases, with a significant linear dependence (Table 2). Extrapolated to 0.0 Å resolution, the ADP estimated errors statistically vanish, as expected for a totally defined system characterized by a near-infinite ratio between number of observations and refined

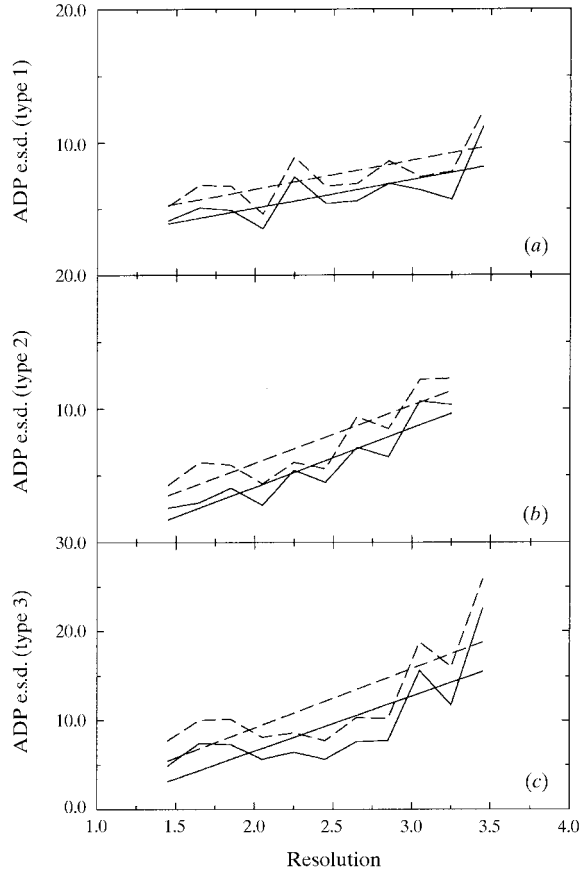

Figure 2
Dependence on crystallographic resolution of the estimated ADP standard errors (e.s.d.) for side-chain (dashed lines) and main-chain (continous lines) atoms. Fitted linear-regression lines are superposed on the observed values. The ADP standard errors are estimated for (*a*) type 1, (*b*) type 2 and (*c*) type 3 sequence pairs.

parameters. The estimated ADP standard errors are larger for side-chain atoms than for main-chain atoms and are larger for solvent-exposed atoms than for those buried in the protein core, an expected result given the ease with which buried and main-chain atoms can be located in electron-density maps. Surprisingly, the differences between estimated ADP errors of main-chain and side-chain atoms and between surface and buried atoms do not depend on resolution. This contrasts with the observation that, at very low resolution, the ADPs normalized to zero mean and unit variance for main-chain and side-chain atoms are identical, while at high resolution they are clearly distinguished (Carugo & Argos, 1997), suggesting that ADPs with similar numerical values may be associated with different standard errors. The need for some estimate of ADP reliability is thus emphasized, especially when atomic positional standard errors are based upon them (Murshudov & Dodson, 1997; Cruickshank, 1996).

The estimated ADP standard errors tend to be larger for type 3 structure pairs, and the gradients with which they increase with lowered resolution tend to decrease in the order type 3, type 2 and type 1. For example, main-chain atom ADPs at 2.0 Å (and 3.0 Å) resolution have estimated errors, based on the regression analyses of Table 2, of 4.9 (7.0), 4.1 (8.5) and 6.6 (12.8) Å^2 if type 1, 2 and 3 structure pairs are used, respectively. Since type 1 and 2 pairs are copies of the same polypeptide polymer in the same protein structure (but independently refined in the case of type 2 pairs), the differences in estimated ADP standard errors reflect different refinement procedures and perhaps experimental conditions (pH, solvent, ionic strength *etc.*). In contrast, type 3 pairs are copies of the same polypeptide polymer in different protein structures with different space groups, and the estimated ADP standard errors thus reflect important physical differences that may, for example, involve significant changes in protein flexibility where different fractions of the cell volume are occupied by solvent and crystal-packing constraints are different. The data presented here suggest that different refinement procedures (different uses of the same algorithm as well as use of different algorithms) have a minor impact on the ADP reliability relative to true physical changes in the crystal architecture, and the complex flexibility of proteins is thus at least qualitatively monitored by the experimentally derived ADPs, mitigating the need for normalization when experimental conditions are similar. Nevertheless, the dependence on resolution of the ADP standard errors estimated with type 1 and 2 pairs is not always similar, especially at very low resolution, where the estimated ADP standard errors based on type 2 pairs is larger than those resulting from type 1 pairs and the final refined parameters are dependent on the refinement strategy. The ADP standard errors estimated with the type 2 pairs are certainly more reliable than those from type 1 and 3 pairs, since the type 2 pairs are physically less different than the type 3 pairs and are independently structurally characterized, in contrast to the type 1 pairs. At 2.0 Å resolution, a value frequently observed in the set of known protein structures examined here, the ADP estimated standard errors are 4.1, 5.9, 5.7, and 3.8 Å^2 for main-chain, side-chain, solvent-exposed and buried atoms, respectively. These errors are large relative to the actual mean ADP values generally found in protein crystal structures (15–20 Å^2) and suggest that very high resolution is needed to accurately monitor protein flexibility with ADPs.

Table 2
Linear-regression analysis of the dependence of the ADP estimated standard error (σ) on the crystallographic resolution (Res) ($\sigma = b_1 + b_2$Res).

Type of structure pairs	Type of atoms	Intercept b_1 (e.s.d.)	Slope b_2 (e.s.d.)	Linear correlation coefficient
Type 1	Main chain	0.7 (1.8)	2.1 (0.7)	0.693
Type 1	Side chain	2.1 (1.9)	2.2 (0.7)	0.746
Type 1	Solvent exposed	1.3 (1.2)	2.4 (0.5)	0.866
Type 1	Protein core	−1.4 (1.4)	2.9 (0.5)	0.869
Type 2	Main chain	−4.7 (1.7)	4.4 (0.7)	0.916
Type 2	Side chain	−2.7 (2.0)	4.3 (0.8)	0.840
Type 2	Solvent exposed	−4.3 (1.8)	5.0 (0.7)	0.921
Type 2	Protein core	−5.2 (1.9)	4.5 (0.8)	0.901
Type 3	Main chain	−5.8 (4.4)	6.2 (1.8)	0.761
Type 3	Side chain	−4.2 (4.7)	6.7 (1.8)	0.768
Type 3	Solvent exposed	−4.8 (5.0)	6.6 (2.0)	0.741
Type 3	Protein core	−7.0 (3.6)	6.1 (1.4)	0.815

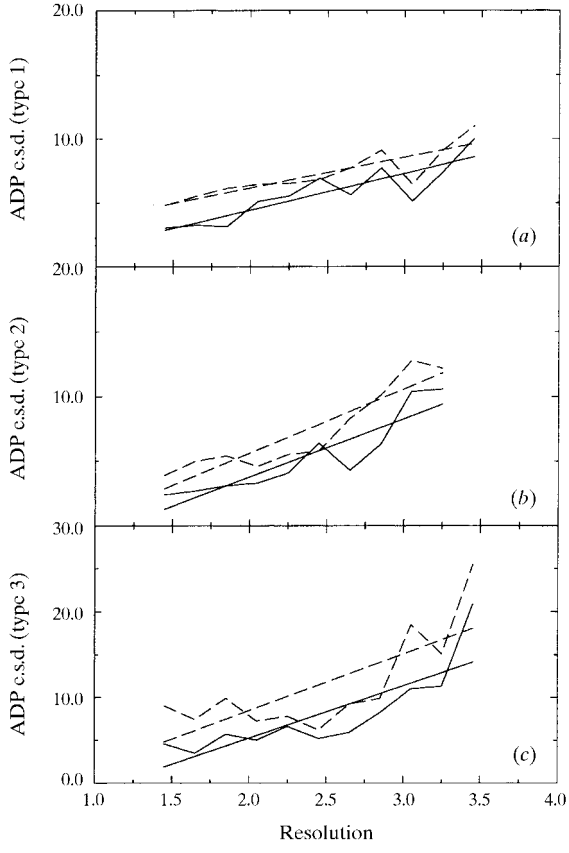

Figure 3
Dependence on crystallographic resolution of the estimated ADP standard errors (e.s.d.) for solvent-exposed (dashed lines) and buried (continous lines) atoms. Fitted linear-regression lines are superposed on the observed values. The ADP standard errors are estimated from (*a*) type 1, (*b*) type 2 and (*c*) type 3 sequence pairs.

Table 3
Mean values of the differences (Å^2) between estimated ADP standard errors for solvent-exposed and buried atoms in various secondary-structural conformations (H, helix; B, strand; T, turn and C, coil).

Secondary structures	Type of structure pairs	Main-chain buried	Main-chain exposed	Side-chain buried	Side-chain exposed
H − B	Type 1	1.4	0.4	2.7	0.2
	Type 2	1.2	0.7	3.1	1.2
	Type 3	1.7	0.7	3.9	2.2
H − T	Type 1	0.8	1.7	2.4	1.5
	Type 2	1.1	1.9	3.0	1.4
	Type 3	0.9	3.5	3.4	3.8
H − C	Type 1	1.0	1.3	2.6	0.8
	Type 2	1.2	0.9	2.9	0.4
	Type 3	1.2	2.5	3.2	1.7
B − T	Type 1	−0.6	−2.0	−0.3	−1.6
	Type 2	−0.2	−2.6	−0.1	−2.7
	Type 3	−0.9	−3.8	−0.5	−5.5
B − C	Type 1	−0.5	−1.8	−0.2	−0.8
	Type 2	−0.2	−1.6	−0.3	−1.7
	Type 3	−0.6	−3.0	−0.6	−3.6
T − C	Type 1	0.1	0.4	−0.2	0.7
	Type 2	−0.1	1.0	−0.2	1.1
	Type 3	0.3	0.8	−0.1	1.9

The linear relationship between estimated ADP standard error and resolution is surprising. It would be expected that the ADP standard error would vary proportionally with the ratio of the number of parameters to be refined to the number of experimental observations. Assuming four variables (x, y, z and the isotropic ADP) for each protein non-H atom which occupies a volume of about 20 Å^3 and disregarding the refinement restraints, the ratio is $(3V_p\mathrm{Res}^3)/(20\pi)$ where Res is the crystallographic resolution and V_p is the fraction of the unit-cell volume occupied by the protein (see Fig. 4). Only for type 3 structure pairs does an approximate relationship exist (Figs. 2 and 3), while for types 1 and 2 the dependence is clearly different, as evidenced by insignificant improvements in correlations using a cubic relationship (in lieu of the linear fit), as shown in Figs. 2 and 3.

The lack of cubic dependence between estimated ADP standard errors and resolution points to the weight of the refinement restraints in determining experimentally derived ADPs. If this weight is assumed to be zero for a certain resolution range, it can be estimated at another lower resolution. For example, if the main-chain atom ADP standard errors derived from the type 1 structure pairs are unaffected by the refinement restraints at 0.7 and 0.8 Å resolution, coefficients a and b of the function $\sigma = a + b\mathrm{Res}^3$ can be determined, where σ is the ADP standard error estimated by the linear function with coefficients given in Table 2 at 0.7 and 0.8 Å resolution and Res is the resolution. At 2.0 Å resolution, the ADP standard error estimated by the cubic function would be 11.7 Å^2, 2.4 times larger than that given by the linear function (4.9 Å^2) with coefficients reported in Table 2. At 3.0 Å resolution it would be 35.2 Å^2, five times larger than that resulting from the linear function (7.0 Å^2). These results clearly point to an underestimation of the ADP standard errors relative to those at 0.7 and 0.8 Å. Although these estimates do depend on the resolution range in which standard errors are assumed to be unaffected by refinement restraints, a useful relative estimation is nonetheless possible.

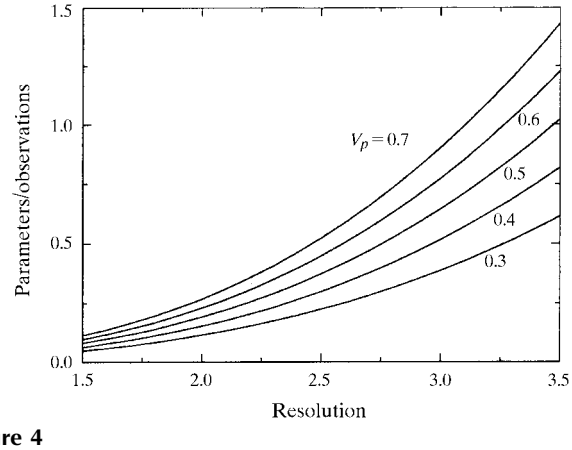

Figure 4
Dependence of crystallographic resolution on the ratio of the numbers of parameters refined to the number of diffraction data. The various curves are computed for different values of the fraction of the unit-cell volume occupied by the protein (V_p).

The estimated standard errors cannot be reliably evaluated for ADPs of atoms in residues with a given secondary-structural conformation, owing to insufficient data in several single-structure pairs (Hamilton, 1972). Nevertheless, some statistically significant trends can be found by comparing all pairs of secondary-structural types (Table 3) over the various structure pairs where there are at least 30 residues with the relevant secondary structure in each protein structure. The differences between estimated ADP standard errors of various secondary-structural types do not depend on resolution. Surprisingly, the estimated ADP standard errors of atoms in helical residues tend to be larger. Buried atoms in turn and coil residues have similar estimated ADP standard errors, while surface atoms in coil residues have smaller ADP standard errors than atoms in turns. The estimated ADP standard errors are smallest in strand residues. Surprisingly, atoms in helices have less reliable ADPs, especially for buried atoms, despite their well known tendency to be less mobile. This may result from inaccuracy in the geometrical or ADP refinement restraints. Some geometrical restraints could be less appropriate for helical moieties relative to other backbone conformations, such as assumptions on the planarity of the peptide groups. The local network of strong hydrogen bonds within helices might allow larger deviations from planarity than in other secondary-structural types. Further, ADP refinement restraints may result in helix-ADP systematic errors owing to the extended local network of short non-bonding contacts, which could excessively smooth the local ADP variability.

4. Conclusions

An estimation of the ADP standard errors has been performed by normal probability plot analyses of pairs of known tertiary protein structures with identical sequences. At the typical 2.0 Å resolution, the ADPs are determined with a reliability of about 5 $Å^2$, a high value compared with the average 15–20 $Å^2$ ADP typically found in protein structures. The errors are also about one order of magnitude larger than those of small asymmetric unit crystal structures. Though the ADP reliability increases with improved crystallographic resolution, it is much less than would be expected on the basis of the increased diffraction data. This allows empirical corrections of estimated ADP standard errors and comparisons of structural data at different resolutions. Estimated ADP standard errors tend to be smaller for main-chain atoms relative to side-chain atoms and for buried atoms relative to solvent-exposed atoms. Estimated ADP standard errors are larger for atoms found in helical residues, suggesting that refinement restraints should be made dependent on backbone conformation.

These results should prove useful in any application of ADPs, such as the analysis and prediction of protein flexibility and the assessment of protein crystal structure quality. The estimation of ADP standard errors makes unnecessary any preliminary data normalization with subsequent loss of information and allows other more sophisticated data analyses.

Financial support from MURST and the kind hospitality of Matti Saraste at the EMBL are gratefully acknowledged.

References

Abrahams, S. C. & Keve, E. T. (1971). *Acta Cryst.* A**27**, 157–165.
Bernstein, F. C., Koetzle, T. F., Williams, G. J. B., Meyer, E. F. Jr, Brice, M. D., Rodgers, J. R., Kennard, O., Shimanouchi, T. & Tasumi, M. (1977). *J. Mol. Biol.* **112**, 535–542.
Bhaskaran, R. & Ponnuswamy, P. K. (1988). *Int. J. Pept. Protein Res.* **32**, 241–255.
Burling, F. T. & Brünger, A. T. (1994). *Isr. J. Chem.* **34**, 165–175.
Carugo, O. & Argos, P. (1997). *Protein Eng.* **10**, 777–787.
Carugo, O. & Argos, P. (1998). *Proteins Struct. Funct. Genet.* **31**, 201–213.
Carugo, O. & Eisenhaber, F. (1997). *J. Appl. Cryst.* **30**, 547–549.
Cruickshank, D. W. J. (1996). *Proceedings of the CCP4 Study Weekend*, edited by E. Dodson, M. Moore, A. Ralph & S. Bailey, pp. 11–22. Warrington: Daresbury Laboratory.
Dauter, Z., Lamzin, V. S. & Wilson, K. S. (1995). *Curr. Opin. Struct. Biol.* **5**, 784–790.
Eisenhaber, F., Lijnzaad, P., Argos, P., Sander, C. & Scharf, M. (1995). *J. Comput. Chem.* **16**, 273–284.
Frauenfelder, H. & Petsko, G. A. (1980). *Biophys. J.* **32**, 465–478.
Frishman, D. & Argos, P. (1995). *Proteins*, **23**, 566–579.
Hamilton, W. C. (1972). *Acta Cryst.* A**28**, 215–218.
Harata, K., Abe, Y. & Muraki, M. (1998). *Proteins Struct. Funct. Genet.* **30**, 232–243.
Heringa, J., Sommerfeldt, H., Higgins, D. & Argos, P. (1992). *CABIOS*, **8**, 599–600.
Hubbard, S. J. & Argos, P. (1995). *Protein Eng.* **8**, 1011–1015.
Jia, Z., Hasnain, S., Hirana, T., Lee, X., Mort, J. S., To, R. & Huber, C. P. (1995). *J. Biol. Chem.* **270**, 5527–5534.
Karplus, P. A. & Schulz, G. E. (1985). *Naturwissenschaften*, **72**, 212–213.
Longhi, S., Czjzek, M., Lamzin, V., Nicolas, A. & Cambillau, C. (1997). *J. Mol. Biol.* **268**, 779–799.
Luedemann, S. K., Carugo, O. & Wade, R. K. (1997). *J. Mol. Model.* **3**, 369–374.
Murshudov, G. N. & Dodson, E. J. (1997). *CCP4 Newslett. Protein Crystallogr.* **33**, 31–39.
Parthasarathy, S. & Murthy, M. R. N. (1997). *Proteins Struct. Funct. Genet.* **6**, 2561–2567.
Parthasarathy, S. & Murthy. M. R. N. (1998). *Proteins Struct. Funct. Genet.* **7**, 525–525.
Peters-Libeu, C. & Adman, E. T. (1997). *Acta Cryst.* D**53**, 56–77.
Ringe, D. & Petsko, G. A. (1986). *Methods Enzymol.* **131**, 389–433.
Vihinen, M., Torkkila, E. & Riikonen, P. (1994). *Proteins Struct. Funct. Genet.* **19**, 141–149.